NIST Big Data Working Group (NBD-WD)
NBD-WD-2013/N0003

Source: **NBD-WG**

Status: **Approved**

Title: **NBD-WG Teleconference Minutes for July 3, 2013, 13:00 – 15:00 EDT**

Author: **NBD-WG Co-Chairs**

The meeting followed the Agenda

I. **Action Items from last meeting**
    a. Send initial subgroup descriptions – Wo
    b. Post Cloud documents – Wo
    c. Provide common tools (document submission, audio/video conf.) – Wo
    d. Reminder: Scope should be clear and shared across subgroups

Wo Chang demonstrated some of the Web tools for Upload document and show Input Document listing which included the currently posted documents for NIST Cloud publications. The Cloud publications will use similar model/approach for our Big Data deliverables which at least include definitions, taxonomies, security reference architecture, reference architecture, and technology roadmap.

One important aspect of document sharing is the document management. For NBD-WG/ Subgroups, all incoming contribution documents will be tracked by using Mxxxx numbering system and Nxxxx for approved output documents. This helps reference any input or output documents as our publication development progress. There is the input document template (http://bigdatawg.nist.gov/Mxxxx_template.docx) available for use. We will be using Microsoft Word as document format for both input and output documents which can enable us to add comments and track changes. The Input Document listing (http://bigdatawg.nist.gov/show_InputDoc2.php) also provides the capability of submitting revision for the same uploaded document.

II. **Welcome Subgroup Co-Chairs**
    a. Five subgroups; three Co-Chairs per subgroup
    b. Subgroups Charter and deliverables (draft)
    c. Responsibility
    d. Meeting schedules
        i.   WG schedules
        ii. Subgroups schedules
        iii. Workshop#1: Save-the-date: September 30, NIST

The NBD-WG Co-Chairs appreciates all those who volunteered to be subgroup Co-Chairs. Due to an overwhelming number of responses, we regretfully were only able to select a limited amount of Co-Chairs for each subgroup. We thank those who responded to our call for your enthusiasm and great support.

The NBD-WG Co-Chairs thank the following members in advance for leading the respective subgroups (Note: The "*" denotes the Lead):

1. **Definitions and Taxonomies**
   *Nancy Grady (SAIC), Natasha Balac(SDSC), Eugene Luster (R2AD)
2. **Requirements**
   *Geoffrey Fox (U. of Indiana), Joe Paiva (VA), Tsegereda Beyene (Cisco)
3. **Security and Privacy**
   *Arnab Roy (CSA, Fujitsu), Nancy Landreville (U. of MD), Akhil Manchanda (GE)
4. **Reference Architecture**
   *Orit Levin (Microsoft), James Ketner (AT&T), Don Krapohl (Augmented Intelligence)
5. **Technology Roadmap**
   *Carl Buffington (USDA), Dan McClary (Oracle), David Boyd (Data Tactic)

Formation of the Subgroup: There will be one Lead Co-Chair, two Co-Chairs:
- Lead Co-Chair – sends meeting agenda with consensus from other Co-Chairs, calls for meeting, conducts meeting, and acts as the main point of contact for the subgroup
- 2 Co-Chairs – help create approach to meet NBD-WG needs and carryout subgroup tasks.

NBD-WG will provide a draft charter for each subgroup with the scope statement and deliverables which each subgroup can refine before presenting to their subgroup members. There will be a teleconference for subgroup Co-Chairs on Monday, July 8 from 2:00 - 3:00 PM EDT. This meeting will discuss video training, charter, and deliverables,

III. **Overall Big Data Approach – Standard Perspective (as a suggestion)**
   a. How can five subgroups achieve common goals?
      i. Definitions and Taxonomies Subgroup
      ii. Security and Privacy Subgroup
      iii. Requirements Subgroup
      iv. Reference Architecture Subgroup
      v. Technology Roadmap Subgroup
   b. How can industry, academia, SDOs, and government achieve common goals?

Wo Chang gave a presentation as a personal suggestion (not a requirement) of the Workgroup Process and Goals. His slides will be circulated to the Working Group. The main points for NBD-WG/Subgroups are to think outside-the-box and create reference architecture similar to existing and successful standards or best practices (PCI card, browser plug-ins, Internet packets, HTML5, etc.):
- Focus using a technology, platform, and infrastructure neutral/agnostic approach
- Enable integration of legacy technologies and future solutions with capability and flexibility for Big Data stakeholders to pick-and-choose best technology while allowing value-added from the Big Data service providers.
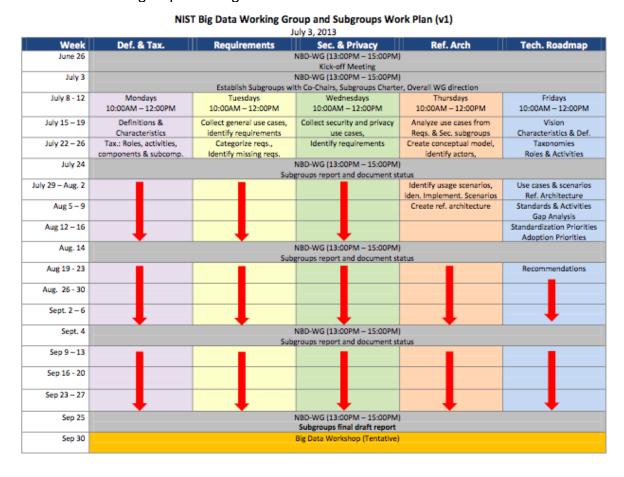
IV. **Contributors to NBD-WG**
   a. Contributions to subgroups via proposal submission
   b. NBD-WG and Subgroups Co-Chairs will identify co-authors and reviewers for NIST Special Publications

All incoming proposal documents may potentially be part of the NBD-WG publications. NBD-WG strongly encourages members to actively contribute ideas and participate in each subgroup email discussion. NBD-WG/Subgroups will identify active members as potential co-authors and reviewers for our publication development.

V.   Q/A

See the edited Chat Session at the end of the Minutes.

Summary: It was decided that a concise summary of teleconference discussion and action item would be added to the Minutes. The key action items for next week are a meeting for all subgroup Chairs on Monday July 8 from 2:00 - 3:00 PM EDT to discuss subgroup charters and deliverables. Tentatively we are proposing 10:00AM – 12:00PM EDT to kickoff subgroup meetings next week. See the Work Plan below.

NIST Big Data Working Group and Subgroups Work Plan (v1)
July 3, 2013

Week	Def. & Tax.	Requirements	Sec. & Privacy	Ref. Arch	Tech. Roadmap
June 26	NBD-WG (13:00PM – 15:00PM) Kick-off Meeting				
July 3	NBD-WG (13:00PM – 15:00PM) Establish Subgroups with Co-Chairs, Subgroups Charter, Overall WG direction				
July 8 - 12	Mondays 10:00AM – 12:00PM	Tuesdays 10:00AM – 12:00PM	Wednesdays 10:00AM – 12:00PM	Thursdays 10:00AM – 12:00PM	Fridays 10:00AM – 12:00PM
July 15 – 19	Definitions & Characteristics	Collect general use cases, identify requirements	Collect security and privacy use cases,	Analyze use cases from Reqs. & Sec. subgroups	Vision Characteristics & Def.
July 22 – 26	Tax.: Roles, activities, components & subcomp.	Categorize reqs., Identify missing reqs.	Identify requirements	Create conceptual model, identify actors,	Taxonomies Roles & Activities
July 24	NBD-WG (13:00PM – 15:00PM) Subgroups report and document status				
July 29 – Aug. 2				Identify usage scenarios, iden. Implement. Scenarios	Use cases & scenarios Ref. Architecture
Aug 5 – 9				Create ref. architecture	Standards & Activities Gap Analysis
Aug 12 – 16					Standardization Priorities Adoption Priorities
Aug. 14	NBD-WG (13:00PM – 15:00PM) Subgroups report and document status				
Aug 19 - 23					Recommendations
Aug. 26 - 30					
Sept. 2 – 6					
Sept. 4	NBD-WG (13:00PM – 15:00PM) Subgroups report and document status				
Sep 9 – 13					
Sep 16 - 20					
Sep 23 – 27					
Sep 25	NBD-WG (13:00PM – 15:00PM) Subgroups final draft report				
Sep 30	Big Data Workshop (Tentative)				

Edited Chat Room Discussion from July 3 Meeting

(10:12 AM) Karen Guertler: Suggest that the subteams publish content in Word format, to allow group members to provide feedback via markup edits. Alternatively, we can add *comments* via PDF... but not readily markup the text

(10:13 AM) Karen Guertler: Thanks - that makes sense, as the Cloud group owns that content. :)
(10:31 AM) Felix Njeh - COMINT: Big Data in the Cloud sounds like a good approach
(10:33 AM) Gregg Brown: You can leave us off, we're used to it, :-) but no EC2?
(10:33 AM) Karen Guertler: OK - in reference to the current slide - I expect that the models / reference architectures will be vendor neutral.
(10:34 AM) Felix Njeh - COMINT: I agree - the ref model should be vendor agnostic
(10:39 AM) Keith Hare, JCC Consulting, Inc.: May also want to look at the SQL standards
(10:41 AM) Keith Hare, JCC Consulting, Inc.: It would be useful to have this presentation for offline review.
(10:42 AM) Ron Hale ISACA: It seems that the key to success will be to have shared use cases between the groups
(10:42 AM) Karen Guertler: Ron makes an excellent point.
(10:42 AM) Ron Hale ISACA: We should be general enough so that the solutions are not specific to cloud but also include private implementations
(10:42 AM) Felix Njeh - COMINT: I have a draft ref model similar to Wo's I can work with him and maybe by next meeting we can share???
(10:44 AM) Karen Guertler: And @Ron, it seems that different folks have a different perspective re the definition of a use case.
(10:44 AM) Ashit: I agree that big data ref frameworks should not be solely tied to the cloud - we should allow private implementations
(10:45 AM) Ron Hale ISACA: If there are other definitions of use cases it will be important to come up with a consistent definition before we get to far into the project
(10:45 AM) Gary Mazzaferro: I have a RA also similar... Presented about a year ago to DISA, ONI intel communities. I cna scrub them fot general public
(10:46 AM) Keith Hare, JCC Consulting, Inc.: @Felix - I would like to see multiple models from different points of views.
(10:50 AM) Vivek Navale: since there will be considerable diversity in use cases, depending on disciplines, so with the time lines we have for the September meeting, do we have subject areas mapped out for use cases?.
(10:53 AM) Tim Zimmerlin (Automation Technologies): Will the NBD-WG be identifying existing standards and their standards groups?
(10:53 AM) Ron Hale ISACA: Another challenge in use cases is to anticipate future needs and directions. Did anyone have any idea where the web would be 10 years after it was first released for public use
(10:54 AM) Gregg Brown: The Cloud Work had an organizing principle of helping US Fed Agencies to use the cloud. Do we have a similar focus for Big Data?
(10:55 AM) Keith Hare, JCC Consulting, Inc.: @Tim -- as a participant in the US and International SQL Standards process, I expect to articulate what pieces already exist in the SQL Standards..
(10:58 AM) Chaitan Baru (SDSC): i think it would be feasible, and probably quite interesting, to make the definitions primarily based on the "state" of the data itself, rather than based on technologies.
(11:00 AM) Vivek Navale: Would it be useful to select use case from areas such as Health, Energy, Space, Commerce, Defence etc.

(11:02 AM) Orit Levin (Microsoft): I share the Chaitan Baru (SDSC) view.

(11:02 AM) Ashit: Domain-specific use-cases is one approach. Another would be categorizing them based on data modalities and streaming, size, distributed data, etc.

(11:04 AM) Gary Mazzaferro: Its good to summarize the meeting minutes at mid points and end

(11:07 AM) Karen Guertler: Completely agree - and defining next steps.

(11:09 AM) Ron Hale ISACA: Part of the agenda should be questions that each of the subgroups want to bring to the larger group. This will help to increase consistent thinking and support the need for groups to test concepts in a larger group.

(11:10 AM) Bob Marcus (ET-Strategies): Due to the tight schedule, I think the goal for September should be for subgroups to determine and describe consensus agreements and document areas of disagreement and propose alternatives

(11:13 AM) Arnab Roy (Fujitsu): What is the timeline for drafting the subgroup charters?

(11:14 AM) Gary Mazzaferro: I'd like to gain a consensus on purpose and stakeholder priority

(11:15 AM) Quyen Nguyen: Seems like Ref Arch. depends on Use Cases from Req and Sec groups. How can Ref Arch start early then?

(11:18 AM) Bob Marcus (ET-Strategies): What are action items?

(11:18 AM) Gary Mazzaferro: Part2: as some base assumptions to move fwd

(11:20 AM) Karen Guertler: Just agreeing with Gary & Quyen's observations.

(11:20 AM) Bob Marcus (ET-Strategies): It is critical that subgroup decide on their deliverables for September. Everything else should be postponed until after September 30.

(11:22 AM) Vivek Navale: I would add that if there are dependancies between the different sub-groups activities, they need to identified quickly as well.

(11:22 AM) Chaitan Baru (SDSC): Just checking: did we say that there will be a huddle of the subgroup co-chairs in order to get on same page?